

WHAT IS CLAIMED IS:

1 1. A method of analyzing gene expression, gene annotation, and sample information
2 in a relational format supporting efficient exploration and analysis, the method comprising:

3 providing a data warehouse which comprises a gene expression database for storing
4 quantitative gene expression measurements for tissues and cell lines screened using various
5 assays; a clinical database for storing information on bio-samples and donors; and a fragment
6 index for biological properties for DNA fragments;

7 receiving a query regarding gene expression of one or more DNA fragments; determining
8 the level of gene expression of the one or more DNA fragments; correlating the level of
9 gene expression with the clinical database and the fragment index; and
10 displaying the results of said correlation.

1 2. The method of claim 1, wherein the data warehouse is constructed in a star
2 relational schema.

1 3. The method of claim 1, wherein the data warehouse is constructed in a
2 snowflake relational schema.

1 4. The method of claim 1, wherein the analysis of gene expression, gene annotation,
2 and sample information further comprises identifying two sets of DNA
3 fragments: those that are consistently expressed within the sample set, and those that are
4 consistently not expressed.

1 5. The method of claim 1, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises a gene signature differential
3 analysis which compares two gene expression signature and derives four sets of DNA gene
4 fragments: those that are in both the first gene signature's present gene set and the second's

5 absent gene set, those that are in both the first gene signature's absent gene set
6 and the second's present gene set, those that are in both present gene sets, those that are in
7 both absent gene sets.

1 6. The method of claim 1, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises a fold change analysis which
3 quantifies the change in expression for differentially expressed genes between pairs of
4 DNA fragments.

1 7. The method of claim 1, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises an E Northern analysis which identifies
3 DNA fragments with regard to a pair of user-selected percentiles over the values for a sample.

1 8. A computer system comprising
2 a data warehouse which comprises a gene expression database for storing quantitative
3 gene expression measurements for tissues and cell lines screened using various assays; a clinical
4 database for storing information on bio-samples and donors; and a fragment index for biological
5 properties for DNA fragments and

6 a user interface capable of receiving a query regarding gene expression of one or more
7 DNA fragments and displaying the results of a correlation of the level of gene expression with
8 the clinical database and the fragment index.

1 9. The computer of claim 8, wherein the data warehouse is constructed in a star
2 relational schema.

1 10. The computer of claim 8, wherein the data warehouse is constructed in a
2 snowflake relational schema.

1 11. The computer of claim 8, wherein the analysis of gene expression, gene

2 annotation, and sample information further comprises identifying two sets of DNA fragments:
3 those that are consistently expressed within the sample set, and those that are consistently not
4 expressed.

1 12. The computer of claim 8, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises a gene signature differential analysis
3 which compares two gene expression signature and derives four sets of DNA gene fragments:
4 those that are in both the first gene signature's present gene set and the second's absent gene set,
5 those that are in both the first gene signature's absent gene set and the second's present gene set,
6 those that are in both present gene sets, those that are in both absent gene sets.

1 13. The computer of claim 8, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises a fold change analysis which quantifies the
3 change in expression for differentially expressed genes between pairs of DNA fragments.

1 14. The computer of claim 8, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises an E Northern analysis which identifies
3 DNA fragments with regard to a pair of user-selected percentiles over the values for a sample.

1 15. A computer program product comprising a computer-usable medium having
2 computer-readable program code embodied thereon relating to a data warehouse which
3 comprises a gene expression database for storing quantitative gene expression measurements for
4 tissues and cell lines screened using various assays; a clinical database for storing information on
5 bio-samples and donors; and a fragment index for biological properties for DNA fragments;
6 the computer program product comprising computer-readable program code for effecting
7 the following steps within a computing system:
8 providing an interface for receiving a query regarding gene expression of one or more

9 DNA fragments;
10 determining the level of gene expression of the one or more DNA fragments;
11 correlating the level of gene expression with the clinical database and the fragment S
12 index; and
13 displaying the results of said correlation.

1 16. The computer program product of claim 15, wherein the data warehouse is
2 constructed in a star relational schema.

1 17. The computer program product of claim 15, wherein the data warehouse is
2 constructed in a snowflake relational schema.

1 18. The computer program product of claim 15, wherein the analysis of gene
2 expression, gene annotation, and sample information further comprises identifying two sets of
3 DNA fragments: those that are consistently expressed within the sample set, and those that are
4 consistently not expressed.

1 19. The method of claim 15, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises a gene signature differential analysis
3 which compares two gene expression signature and derives four sets of DNA gene fragments:
4 those that are in both the first gene signature's present gene set and the second's absent gene set,
5 those that are in both the first gene signature's absent gene set and the second's present gene set,
6 those that are in both present gene sets, those that are in both absent gene sets.

1 20. The computer program product of claim 15, wherein the analysis of gene
2 expression, gene annotation, and sample information further comprises a fold change analysis
3 which quantifies the change in expression for differentially expressed genes between pairs of
4 DNA fragments.

1 21. The method of claim 15, wherein the analysis of gene expression, gene
2 annotation, and sample information further comprises an E Northern analysis which identifies
3 DNA fragments with regard to a pair of user-selected percentiles over the values for a sample.

1

11/11/2011 11:11:11 AM